

# Exhibit C



SIDLEY AUSTIN LLP  
1501 K STREET, N.W.  
WASHINGTON, D.C. 20005  
+1 202 736 8000  
+1 202 736 8711 FAX

AMERICA • ASIA PACIFIC • EUROPE

+1 202 736 8396  
RKEELING@SIDLEY.COM

August 30, 2024

CONFIDENTIAL

**By Email**

Shira A. Steinberg  
Bureau of Competition Attorney  
(202) 326-3633  
ssteinberg1@ftc.gov

Re: *Federal Trade Commission, et al. v. Amazon.com, Inc.*, No. 2:23-cv-01495-JHC

Dear Shira:

I write in response to Plaintiffs' August 13, 2024 letter regarding the parties' respective search terms and TAR 2.0 ("TAR") proposals. Before turning to specifics, however, we want to express a few concerns regarding the negotiations to date and offer some suggestions for making better progress going forward.

Plaintiffs' approach to Amazon's responses to discovery often appears designed to demand that Amazon provide the maximum amount of information or maximum volume of materials through an exhaustive process that would impose the highest possible burden, and to do so all at once. The ESI Order's opening paragraph advises the parties to "cooperate in facilitating and reasonably limiting discovery requests and responses." Dkt. # 256 at ¶(A)(1). Plaintiffs' approach to date is contrary to this principle. We are flagging this issue now because Plaintiffs have expressed concerns about the pace of discovery, but it is our experience that Plaintiffs' approach has impeded progress and we believe there is an opportunity to make serious progress in the TAR, search terms, and custodial negotiations.

Plaintiffs' most recent letter on Search Terms and TAR (the second received on that subject alone in a matter of weeks) is an example of the problematic dynamic identified above.

First, Plaintiffs state their proposed search terms are "aimed at collecting *substantially all documents that may be responsive* to Plaintiffs'" more than 350 requests for production. Ltr. from S. Steinberg to R. Keeling at 1 (Aug. 13, 2024) (emphasis added). Yet, Plaintiffs' letter questions why Amazon did not immediately provide hit reports for terms that have *no connection* to any of the claims or defenses—or even Amazon's business operations at all—or are facially overbroad, include materials Amazon has never agreed to search for, or seem akin to a fishing expedition despite Plaintiffs' 4+ year investigation. These are terms such as OMG\*, f\*ck, sh\*t, shoot, WTF\*, crap, top, desperate, god, Jesus, and holy. As we discussed on our August 22 call,

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 2

a term like sh\*t would also capture shot, shut, shout, shirt, shift, and any other word that starts with “sh” and ends with “t”; no wonder it brings in millions of documents. The religious terms are strange terms for a government body to be proposing in a case that has nothing to do with religion (also, Jesus is a common name). Plaintiffs further sought hit reports on terms like document\* W/10 AND privileged, writ\* W/10 AND attorney\*, and e-mail W/10 AND legal counsel, all of which would undoubtedly return privileged materials—even those on irrelevant topics. Although Amazon has provided the hit reports, we are concerned that the demand for hit reports was itself part of this fishing expedition. We are also concerned that Plaintiffs’ insistence that *Amazon* must justify the burden or provide specific objections to facially problematic terms puts make-work and form over substantive progress.

Second, Plaintiffs seek further revisions to an uncontroversial TAR protocol that is nearly identical to what the California AG agreed to, even though many of Plaintiffs’ proposals add redundant provisions, define commonly used terms, and attempt to micromanage Amazon’s review process. And Plaintiffs surprisingly suggest that Amazon’s previous TAR agreements with FTC, are somehow “not … a useful guidepost for the use of TAR” in this matter.

Third, the scope of Plaintiffs’ demands extends to engineering demands and technological theories. Several pages of Plaintiffs’ letter provide engineering advice as to how our vendor, Consilio, one of the largest and most respected eDiscovery vendors in the world, can scale its technology to prodigious levels in order to account for the vast volume of documents implicated by Plaintiffs’ demands. But for the reasons noted below, *see* pp. 8-9, *infra*, Plaintiffs’ technological response does not appear to be based on a technological understanding.

Amazon now responds to each of these issues in greater detail below. We do so in a good faith attempt to answer Plaintiffs’ questions and to make progress, including by offering a number of compromises. Amazon continues to believe that both parties will have to be reasonable to efficiently move forward. *See* Dkt. # 246 at ¶ A(1). Amazon will continue to do so, and hopes Plaintiffs will as well.

## Search Terms and TAR

Amazon appreciates Plaintiffs’ continued agreement to using search terms in combination with TAR. We also remain aligned, at least in principle, that the use of search terms and TAR will allow the parties to reach agreement on search terms—and more quickly—than we would otherwise be able to agree to in the absence of TAR. Further, we continue to agree that the most efficient way to address Plaintiffs’ expansive discovery requests is through a “bucketing” approach that considers categories of search terms and custodians according to topic.

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 3

As we explained in our August 8, 2024 letter, however, Plaintiffs' proposed set of nearly 1,100 search strings<sup>1</sup> as applied to Amazon's 54 proposed custodians and limited to Amazon's proposed default time period of January 1, 2018, to September 26, 2023, returned more than 15 million documents (including families). Expanding the date range, correcting some of Plaintiffs' errored terms, and increasing custodians to our now-agreed 59 custodians, increased that number to nearly 20 million documents. Extrapolating to Plaintiffs' counterproposal of 148 custodians would yield an estimated 50 million documents, even without considering Plaintiffs' request for Amazon to search for documents going back ten years to 2014.

Moreover, as also explained in our August 8, 2024 letter, many of the terms with the highest counts are from Plaintiffs' "General" category and have no clear relation to Plaintiffs' RFPs or the claims in Plaintiffs' complaint. We identified three specific, facially problematic categories, each with reference to specific terms, and asked Plaintiffs to reasonably narrow, revise, or delete search terms in the "General" category. These categories were (1) swear words without any connection to the complaint;<sup>2</sup> (2) terms clearly meant to capture documents responsive to Plaintiffs' Request No. 8 regarding [REDACTED];<sup>3</sup> and (3) numerous one-word terms that Plaintiffs previously agreed would be inappropriate.<sup>4</sup> In addition to requesting that Plaintiffs reconsider their "general" search terms, Amazon requested that Plaintiffs break out advertising- and search-related terms and custodians from the "Market Definition" category.

The astronomical volumes implicated by Plaintiffs' search terms—potentially tens of millions of documents—also highlight a difference between how the parties view the purpose of search terms here. Plaintiffs expressly state that Plaintiffs' terms are aimed at "collecting substantially *all* documents that may be responsive to Plaintiffs' discovery requests." *See Ltr. from S. Steinberg to R. Keeling at 1 (Aug. 13, 2024)* (emphasis added). Amazon does not agree that this is the appropriate goal. Instead, the search terms should be crafted to return documents

---

<sup>1</sup> As the ESI Order explains, "A disjunctive combination of multiple words or phrases (e.g., 'computer' or 'system') broadens the search, and thus each word or phrase shall count as a separate search term unless they are variants of the same word." Dkt.# 256 at ¶ C(2)(c). Many of Plaintiffs proposed terms use disjunctive connectors, significantly broadening the scope of the terms. By our count, Plaintiffs' proposed terms contain 280 disjunctive connectors.

<sup>2</sup> For example: "Sh\*t OR Shoot" and "F\*ck OR fck OR Fu\*k OR Motherf\* OR WTF\*."

<sup>3</sup> [REDACTED]

<sup>4</sup> For example: "Google," "Target," "Microsoft," "Walmart," "UPS," "Jet," and "eBay."

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 4

that are *likely* to be relevant to the claims and defenses in the litigation.<sup>5</sup> Terms like “hotly debated,” “jet,” “investigation,” “target,” “Instagram,” “leverage,” and so on could, in theory, appear within responsive documents. So could most other terms found in a dictionary. But there is nothing about the presence of each of those words standing alone that would suggest the document is likely to be relevant. Plaintiffs’ attempt to capture “substantially all” documents that “may” be responsive is leading to plainly disproportionate results. Given the extensive pre-suit investigation that you conducted, we would expect Plaintiffs to be far more targeted in their search queries.

Moreover, it is incorrect for Plaintiffs to claim that use of TAR as a backstop “will address any burden concerns” for Amazon. *See* Ltr. from S. Steinberg to R. Keeling at 1 (Aug. 13, 2024). As addressed below, there are significant technological burdens—likely insurmountable ones—just in applying TAR 2.0 at such a scale. And even if those could be addressed, TAR is not a magic wand. There will still be a significant manual review that will need to be conducted by attorneys. For the sake of example, consider a well-performing TAR 2.0 workflow that was able to achieve the target recall level while excluding 70% of documents from human review. That would be a tremendous savings, indeed. But at the 50 million documents projected for Plaintiffs’ proposed custodians and search terms, that would still leave 15 million documents to be reviewed by human attorneys. Assuming a large team of 50 attorney reviewers, completing first-level review would likely take more than 214 weeks—over four years.<sup>6</sup> Clearly Plaintiffs must narrow the search terms—and custodians and time period—to a more manageable volume.

Lastly, we maintain that Plaintiffs’ proposal to apply all of Plaintiffs’ proposed search terms to the CEOs and current and former S-team members is inappropriate. This would result, for example, in applying fulfillment terms to unrelated S-team members like Colleen Aubrey, who served as SVP, Ad Products and Tech, and now serves as SVP, Amazon Web Services (AWS) Solutions. Such obvious mismatches can be easily avoided by negotiating which S-team custodians will be searched for each of the buckets. Given the relatively small number of current and former S-team members, we do not expect that this will result in significant additional delay or work for the parties. In fact, Amazon has already done so as part of its July 22, 2024 proposal. That proposal included Jeff Bezos as a custodian for almost all buckets, showing that Amazon is open to applying search terms broadly to S-team members to the extent that there is good reason to do so.

---

<sup>5</sup> See Fed. R. Civ. P. 26(b)(1).

<sup>6</sup> This assumes a typical review rate of 35 documents/hour at 40 hours/reviewer/week with no holidays.

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 5

## TAR 2.0

We are pleased that the parties made continued progress towards finalizing the TAR Protocol during our August 23, 2024 meet and confer. Indeed, we believe the parties agree or agree in principle on a number of previously-contested issues.

- **Target recall level:** We believe we have agreement in principle on the target recall level that should be applied in this matter. Plaintiffs' proposed a target recall level of 75% that would increase to 85% if either richness or precision were greater than 50%. While such an approach would be fatally problematic if applied to the entire review population, we believe we have an understanding in light of Plaintiffs' explanation in your recent letter and our discussions during the meet and confer.

To recap, Plaintiffs suggested that "if it is more likely than not (i.e., a greater than 50% chance) that Amazon can identify responsive documents in the population, it should continue doing so until approximately 85% of the relevant documents are found in the population." Ltr. from S. Steinberg to R. Keeling at 1 (Aug. 13, 2024) at 7. As we discussed, it is important that the associated metrics be fairly calculated. In a TAR 2.0 workflow, review is prioritized according to each document's score. Near the outset of the review, high scoring documents are reviewed that are quite likely to be responsive, but, as the review continues, lower-scoring documents are reviewed that are far less likely to be responsive. Eventually the review will reach a point of diminishing returns. If the goal is to determine whether there is a 50/50 chance that additional documents will be uncovered from continued review, then that should be determined by the documents reviewed at the time 75% recall is achieved—i.e., the precision observed in the most recent tranche of documents reviewed along the scoring curve. We would accordingly calculate precision by analyzing the number of responsive documents in the 5,000 lowest-scoring documents above the TAR cut-off score at the time 75% recall is achieved. If precision is greater than 50% at the time that 75% recall is achieved, Amazon will continue to review until it either reaches 85% recall or until precision of the lowest-scoring 5,000 documents drops below 50%.

- **Email threading:** As Plaintiffs' note, the parties previously agreed to adopt email threading. We have provided a revised proposal including an email threading provision. Please note that our edit specifies that we will perform email threading on the document population "prior to ingestion *into Brainspace*." See Amazon's Third TAR 2.0 Proposal at ¶ 3.4 (emphasis added). Native documents will still need to be ingested for processing (e.g., metadata and text extraction), as we cannot identify spam without doing so. Any

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 6

documents excluded from the population as a result of email threading will not be included in Brainspace and will not be used to train the TAR 2.0 model.

- **Spam:** Contrary to Plaintiffs' characterization, the parties agree to remove spam and junk documents prior to uploading the document population into Brainspace. *See* Amazon's Second TAR 2.0 Proposal at ¶ 3.5. We addressed the identification and removal of spam and the criteria for doing so by separate letter in advance of finalizing the TAR protocol. *See* Ltr. from R. Keeling to S. Steinberg (Aug. 29, 2024). We have revised Section 3.5 to reflect this.
- **Linear review:** We previously agreed to Plaintiffs' proposal to disclose any custodians or repositories Amazon plans on reviewing in a linear fashion outside of TAR prior to finalizing the TAR protocol. You added a new provision to this effect, which we removed. Section 2.1 of the TAR protocol already included language allowing for the parties to agree in writing that, for certain documents, Amazon will "conduct a linear responsiveness review." We did not see the value in having two provisions address the same issue. Moreover, as yet, we have not identified any custodians or repositories we plan on reviewing in a linear fashion outside of the TAR process.
- **Definitions:** Plaintiffs proposed several definitions of terms commonly used in eDiscovery. We did not think that drafting new definitions of such well known terms was useful. Nor do such definitions need to be in the TAR protocol. Moreover, the definitions provided by Plaintiffs referred to concepts such as uncoded documents above the cutoff score that misunderstand a basic TAR 2.0 approach.<sup>7</sup> If Plaintiffs insist on defining these terms, we can reach agreement on the following definitions:
  - *Recall.* The estimated percentage of responsive documents found, as calculated by the number of documents tagged responsive divided by the total of (1) the number of documents tagged responsive and (2) the estimated number of responsive documents in the un-reviewed portion of the Review Population, expressed as a percentage. The estimated number of responsive documents in the un-reviewed portion of the Review Population is calculated by multiplying the percentage of documents tagged as responsive during the review of a statistical sample drawn from the un-reviewed portion of the

---

<sup>7</sup> Plaintiffs stated that these definitions came from Relativity, but they appear to have been significantly modified. *See* Relativity, Project Validation Statistics, [bit.ly/3AKhLOZ](https://bit.ly/3AKhLOZ).

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 7

Review Population multiplied by the total number of documents in the un-reviewed portion of the Review Population.

- *Richness.* The estimated percentage of documents that are responsive in the Review Population as calculated by: the number of documents coded responsive during the review of a statistical sample drawn from the Review Population divided by the size of the sample, expressed as a percentage.
- *Precision.* The estimated percentage of documents reviewed that were responsive as calculated by: dividing the number of documents tagged responsive by the number of documents reviewed, expressed as a percentage.
- *Elusion Rate.* The estimated percentage of responsive documents in the unreviewed portion of the Review Population as calculated by: dividing the number of documents tagged responsive during the review of a statistical sample drawn from the unreviewed portion of the Review Population divided by the size of the sample, expressed as a percentage.

For other of Plaintiffs' proposed changes to the TAR protocol, Plaintiffs criticize Amazon for deleting them without explanation. Yet Plaintiffs' never explained the basis for their inclusion when they were originally added. Now that Plaintiffs have offered explanations, we respond as follows:

- **Text size:** Excluding over 1MB in extracted text size is far too restrictive and does not accord with modern discovery. Our experience is that Brainspace correctly analyzes larger documents, and we note that the same page cited by Plaintiffs states that Brainspace "no longer excludes files larger than 16MB and is able to ingest most text-based documents with minimal issues." *Id.* Accordingly, we will use a 5MB limit for extracted text size for spreadsheets and a 16MB limit for extracted text size for other files. For comparison, Relativity Active Learning, another popular TAR 2.0 application, allows extracted text files up to 32MB.
- **Bulk coding:** Plaintiffs proposed that Amazon disclose each instance in which it makes use of bulk coding features or disavow the use of bulk coding. We cannot agree to this and have never encountered such a request. This is another example of an attempt to micromanage day-to-day document review—here by limiting or precluding use of a common feature with numerous legitimate uses. It also ignores (or fails to understand due to lack of experience) the realities of modern production reviews. Bulk coding is necessary to conduct substantial production reviews at multiple levels. For example, bulk coding is needed for coding go-get productions where the documents are known to be

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 8

responsive. It is also necessary to identify documents for production and for tracking. These are just to name a few scenarios where bulk coding is useful. As Ninth Circuit courts have often recognized: “Responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information.” *See, e.g., NuVasive, Inc. v. AlphaTec*, 2019 WL 4934477, at \*2 (S.D. Cal. Oct. 7, 2019) (citing *The Sedona Principles*, 3d ed., 19 SEDONA CONF. J. 1, Principle 6, 118 (2018)); *Brewer v. BNSF Rwy. Co.*, 2018 WL 882812, at \*2 (D. Mont. Feb. 14, 2018) (same).

- **Validation:** Plaintiffs proposed that Amazon produce the validation documents instead of providing access to them in a separate database. We do not understand why Plaintiffs would take issue with the same validation process that was used successfully in the California AG litigation and in many Second Requests from the FTC and Department of Justice.<sup>8</sup> The underlying concern with producing validation documents is simple: the documents at issue are randomly selected from a population that is highly likely to be non-responsive and will be withheld only subject to privilege. Accordingly, this population is likely to include sensitive and proprietary documents unrelated to Plaintiffs’ claims and defenses and is outside Rule 26. The California AG conducted its validation review in another database. The FTC has conducted similar validations reviews of Amazon TAR productions in other databases. The same process should apply here.

Our revised TAR protocol is provided as Appendix A.

Without any cited basis in actual experience, Plaintiffs also attempt to cast doubt on Amazon’s concerns that Brainspace becomes unstable when ingesting more than 9 million documents. We shared these concerns with Plaintiffs in an effort to head off potential technological issues that could impede the efficient and orderly production of documents down the line and potentially break the TAR tool. These concerns are based on our own experience as discovery counsel and shared by Consilio. We supported them with citations to Brainspace’s publicly available documentation and the documentation of comparable tools.

In response, Plaintiffs have relayed claims from a conversation with an unnamed salesperson for Brainspace. For example, Plaintiffs cite to this salesperson to the effect that Brainspace has ingested *billions* of documents without issue. *See* Ltr. from S. Steinberg to R. Keeling at 5 (Aug. 13, 2024) (emphasis in original). Not stopping there, Plaintiffs say that this salesperson claimed, that Brainspace “does not have performance limitations.” *Id.* Such boasting

---

<sup>8</sup> We are also surprised that Plaintiffs do not view Second Requests as a “useful guidepost for the use of TAR in litigation.” Ltr. from S. Steinberg to R. Keeling at 5 (Aug. 13, 2024).

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 9

is without any foundation and cannot be taken at face value. This statement further is directly contrary to years of experience from counsel and thousands of matters where Consilio has used Brainspace. But to confirm, we spoke to one of the inventors of Brainspace, who confirmed that Brainspace has never dealt with billions of documents.

Moreover, while it may technically be possible to put very large volumes of documents into Brainspace, that does not mean that Brainspace performs well at those volumes and will not break. When data volumes are larger, making calculations requires more machine resources and machine time, and the likelihood of technological issues increases. Here, Plaintiffs have suggested that Amazon should create multiple Brainspace instances—each with up to 9 million documents—and use portable models to transfer training from one instance to another. We have substantial experience with using portable TAR models. We have used portable TAR models with Brainspace, and also used portable TAR models with other TAR applications, such as Relativity’s RAL tool, Ankura’s Predict applications, and Lighthouse’s Prism TAR tool. In short, portable models do not perform nearly as well as promised—whether that promise comes from sales representatives or otherwise. We consistently see a significant degradation of model performance when the TAR model is taken from one custodial data set and applied to another. In other words, the model does a worse job at finding what it was originally trained to identify.

We also consulted with Consilio, and there are a number of technological problems with Plaintiffs’ proposal. Having multiple Brainspace instances would increase the amount of data being written back to Relativity, a key performance bottleneck. It would also increase the number of fields in the database, leading to “table bloat” that can slow indexing and otherwise impede performance. It would also create significant administrative overhead and burden to keep the instances in sync throughout an ongoing TAR 2.0 review. More generally, portable models were intended to allow customers to use a model developed in one matter on another—e.g., a model that identifies patient health information that should be redacted prior to production. They were not intended to be a scaling solution in a TAR 2.0 review.

The fact that we are even discussing such novel approaches demonstrates the unprecedented demands of Plaintiffs’ discovery in this case and the untargeted nature of Plaintiffs’ search terms. In this respect, we appreciate Plaintiffs’ suggestion that the parties may ultimately agree on a set of search terms that will result in a volume of documents that will not implicate our concerns regarding ingesting more than 9 million documents into Brainspace. See *id.* at 6-7.

Plaintiffs also expressed some question regarding why we believe it is appropriate to put standalone spreadsheets in TAR. As we previously stated, Consilio has included spreadsheets in thousands of matters over the years without issue, and their standard Brainspace setup includes

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 10

spreadsheets by default. This accords with our own experience on countless TAR matters. Brainspace evaluates spreadsheets just like it would evaluate any other document. Spreadsheets are not special, and the TAR model performs very well on them.

Finally, Plaintiffs raised concerns about how differing date ranges for different categories of documents will impact the model. Although we understand these concerns, we do not anticipate any practical problems. For example, we would run fulfillment terms on fulfillment custodians for the agreed-upon time period. We expect that would return the vast majority of fulfillment documents and that those documents would be coded responsive for purposes of training the model. To the extent there are fulfillment-related documents from outside that time period returned by other search terms that were coded non-responsive solely because of the date, that would be a relatively small fraction compared to the number of fulfillment documents identified as responsive. It is, in other words, unlikely to move the needle.

## **Document Collection**

Regarding the status of custodial collections for our 59 proposed custodians, Amazon collects from ten backend sources as part of its standard custodial collection process—i.e., [REDACTED]

[REDACTED] We have completed this process for each of our 54 proposed custodians, and are processing for the additional 5 Amazon recently agreed to. All 54 had custodial data. If Plaintiffs' question is about sources such as laptops, phones, hardcopy documents, and other, custodian-specific sources that require custodial assistance to collect, those have been or will be identified through Amazon's custodial process.

## **Non-Custodial Data Sources**

Our June 20 disclosures referred to Amazon's response to RFP 12 with respect to its identification of non-custodial data sources. That disclosure listed some 20 data sources and categories of data sources. Plaintiffs nonetheless characterize our disclosure as deficient because it did not identify those that are "likely to contain responsive information" and instead contained a disclaimer that "[i]dentification of sources of information does not indicate that such sources or information are relevant or contain responsive information, nor that they can be collected or produced without undue and disproportionate burden." In an effort to clear up any confusion, we previously responded to Plaintiffs' concerns in our July 30 letter. Ltr. from R. Keeling to S. Steinberg (Jul. 30, 2024). Now, as we agreed in our August 23 meet and confer, please see Appendix B for an updated disclosure.

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 11

In response to Plaintiffs' letter, [REDACTED] are properly considered custodial. Regardless, we are collecting all documents that any of our custodians have access to for these data sources—not merely those where they are the author/owner—consistent with Plaintiffs' desired search strategy for these data sources in Plaintiffs' August 13 letter.<sup>10</sup> As a result, we think there is little practical difference between calling them "custodial" or "non-custodial," and we ask that Plaintiffs consider this for purposes of our disclosures as well, as reflected in Appendix B.

Finally, Plaintiffs have raised questions about two data [REDACTED]

[REDACTED] You expressed interest in obtaining S-team goals. To the extent that S-team members are identified as custodians, we agree to collect goals for each S-team custodian and to search for any relevant, responsive goals to the extent such data exists and can be collected and produced without undue burden. [REDACTED]

[REDACTED] If there are other non-custodial data sources identified in the five years since Plaintiffs' investigation commenced that you believe are likely to contain relevant information and that you have not previously raised with us, please do so.

There may be other non-custodial data sources that will be uncovered during the course of our review and, if so, we will collect from those as appropriate.<sup>11</sup>

## Timing

As we explained at the outset of this letter, it is difficult to reconcile Plaintiffs' demand that Amazon substantially complete document production by the end of year given Plaintiffs'

9 [REDACTED]

<sup>10</sup> Search and collection methods for other custodial and non-custodial data sources are source dependent.

<sup>11</sup> Amazon's disclosures are notably broader than Plaintiffs'. The FTC disclosed only "shared drive folders (K: Drive) associated with investigation file Nos. 191-0129 and 191-013 and the FTC's Freedom of Information Act office (stored in Microsoft Azure Cloud), internal webpages ('intranet') for the FTC, and the non-custodial email inbox associated with the address antitrust@ftc.gov (hosted in Microsoft Office 365)." Minnesota, Rhode Island, and Maryland disclosed *no* non-custodial data sources.

# SIDLEY

*Federal Trade Commission, et al. v. Amazon.com, Inc.*

No. 2:23-cv-01495-JHC

Page 12

approach to these negotiations. Despite having conducted a multi-year investigation and receiving some 1.7 million documents in advance of the complaint, Plaintiffs served 393 requests for production. Plaintiffs seek documents going back more than 10 years. Plaintiffs have requested that we produce documents from nearly 150 custodians, including 95 additions. Plaintiffs have proposed more than 1,000 search strings, which are extremely broad, in an effort to identify “substantially all documents that may be responsive.” And in response to our concerns about the burden and even technological feasibility of accommodating these demands, Plaintiffs have proposed unworkable engineering solutions based on a conversation you had with a salesperson.

At this stage, the parties have not yet agreed on the custodians, time period, and search terms that will largely define the scope of the document production. Because the time needed to complete a document review depends on the number of documents that need to be reviewed, setting a deadline for substantial completion at this stage puts the cart before the horse. Amazon is amenable to discussing interim discovery guidelines once the parties have a sense of the scope of the review.

\* \* \*

As always, Amazon is available to meet and confer as necessary to discuss the issues outlined above.

Sincerely,



Robert D. Keeling  
Partner